6.S891 Lecture 24: The Gurvits Capacity

Kuikui Liu

December 7, 2023

In the second half of the previous lecture, we built a deterministic approximate counter for the number of bases of a matroid with simply exponential multiplicative error. In this lecture, we use similar ideas to approximate the permanent of a nonnegative matrix and the mixed volumes of convex bodies.

1 Deterministically Approximating the Permanent

Recall the permanent of a matrix $A \in \mathbb{C}^{n \times n}$ is defined as

$$\operatorname{per}(A) \stackrel{\mathsf{def}}{=} \sum_{\sigma \in S_n} \prod_{i=1}^{n} A_{i,\sigma(i)}, \tag{1}$$

where we write $S_n \subseteq [n]^n$ for the group of permutations $\sigma : [n] \to [n]$.

Theorem 1.1 ([Gur06])**.** *There is a deterministic polynomial-time algorithm which outputs a e^n-multiplicative approximation to $\operatorname{per}(A)$ for any nonnegative matrix $A \in \mathbb{R}_{\geq 0}^{n \times n}$.*

Remark 1. Note that the seminal work of Jerrum–Sinclair–Vigoda [JSV04] gives an FPRAS for this problem. We also previously saw a quasipolynomial-time deterministic algorithm when A is not too far from the all-ones matrix using Barvinok's polynomial interpolation method.

We prove Theorem 1.1 using entropy and log-concavity ideas. In particular, as we mentioned previously for matroid bases, one way to take advantage of the Donsker–Varadhan representation of KL-divergence

$$\mathscr{D}_{\mathrm{KL}}(\nu \,\|\, \mu) = \sup_{f} \left\{ \mathbb{E}_{x \sim \nu}[f(x)] - \log \mathbb{E}_{x \sim \mu}\left[e^{f(x)}\right] \right\} \tag{2}$$

is to view the measure ν of interest as a *restriction* of a "nicer" measure μ on a possibly enlarged state space. For the permanent, we can take μ to be the product measure $\bigotimes_{i=1}^{n} \mu_i$ over $[n]^n$, where $\mu_i(j) \propto A_{ij}$ for each $j \in [n]$. Clearly, $\mu(\sigma) \propto \prod_{i=1}^{n} A_{i,\sigma(i)}$ for every function $\sigma : [n] \to [n]$, including those which are not permutations. Then ν is the restriction of μ to the set of permutations $S_n \subseteq [n]^n$. This observation (plus some simple rearranging) establishes the following convex programming interpretation of the permanent.

Fact 1.2 (Permanent as a Convex Program)**.** *Let $A \in \mathbb{R}_{\geq 0}^{n \times n}$ be any nonnegative matrix. Let ν be the probability measure over $S_n \subseteq [n]^n$ given by $\nu(\sigma) \propto \prod_{i=1}^{n} A_{i,\sigma(i)}$. Then*

$$\log \operatorname{per}(A) = \inf_{f : [n]^n \to \mathbb{R}} \left\{ \log \left(\sum_{\sigma \in [n]^n} e^{f(\sigma)} \prod_{i=1}^{n} A_{i,\sigma(i)} \right) - \sum_{\sigma \in S_n} \nu(\sigma) f(\sigma) \right\}. \tag{3}$$

The hope is then to find a nice subclass of functions f for which the optimization in Eq. (3) becomes tractable. We want this class of functions to be rich enough so that we do not lose too much in the restriction. At the same time, we want $\sum_{\sigma \in [n]^n} e^{f(\sigma)} \prod_{i=1}^{n} A_{i,\sigma(i)}$ and $\sum_{\sigma \in S_n} \nu(\sigma) f(\sigma)$ to be efficiently computable.

Perhaps a natural choice is for f to be *linear*, i.e. $f(\sigma) = \sum_{i=1}^{n} \boldsymbol{w}_{i,\sigma(i)}$ for some matrix of weights $\boldsymbol{w} \in \mathbb{R}^{n \times n}$. The first term is still easy to compute, but unfortunately, $\sum_{\sigma \in S_n} \nu(\sigma) \sum_{i=1}^{n} \boldsymbol{w}_{i,\sigma(i)}$ still seems nontrivial since it depends on the marginal probabilities $\Pr_{\nu}[\sigma(i) = j]$ of matching i to j for all $i, j \in [n]$. However, if we further restrict $\boldsymbol{w}$ to be of the form $\mathbf{1}\boldsymbol{v}^{\top}$, then $f(\sigma) = \sum_{i=1}^{n} \boldsymbol{v}_{\sigma(i)}$ is *constant* over permutations, and given by $\langle \boldsymbol{v}, \mathbf{1} \rangle$. The following result of Gurvits bounds the approximation error for this nice restriction.

Theorem 1.3 ([Gur06]). *Consider the following restriction of Eq. (3) to the functions f of the form $f(\sigma) = \sum_{i=1}^n \boldsymbol{v}_{\sigma(i)}$ over $[n]^n$, for some $\boldsymbol{v} \in \mathbb{R}^n$:*

$$\mathcal{F}_{\mathsf{Cap}}(A) \stackrel{\mathsf{def}}{=} \inf_{\boldsymbol{v} \in \mathbb{R}^n} \left\{ \log \left(\prod_{i=1}^n \sum_{j=1}^n A_{i,j} e^{\boldsymbol{v}_j} \right) - \sum_{j=1}^n \boldsymbol{v}_j \right\}. \tag{4}$$

Then

$$\log \operatorname{per}(A) \le \mathcal{F}_{\mathsf{Cap}}(A) \le \log \left(\frac{n^n}{n!} \right) + \log \operatorname{per}(A). \tag{5}$$

Let us first discuss how Theorem 1.3 yields an elementary proof of the van der Waerden conjecture on the permanent of *doubly stochastic* matrices. These are nonnegative matrices $A \in \mathbb{R}_{\ge 0}^{n \times n}$ such that its row sums and column sums are all 1. The following corollary gives a flavor of how these types of inequalities can be applied to problems in combinatorics.

Corollary 1.4 (Resolution of van der Waerden's Conjecture; [Gur06]). *For every doubly stochastic matrix $A \in \mathbb{R}_{\ge 0}^{n \times n}$, we have $\frac{n!}{n^n} \le \operatorname{per}(A) \le 1$.*

Note that both bounds are tight; the upper bound is attained by the identity matrix, while the lower bound is attained by the normalized all-ones matrix $\frac{1}{n}\mathbf{1}\mathbf{1}^\top$. This was first established independently by Egorychev and Falikman [Ego81; Fal81].

Proof of Corollary 1.4. By Theorem 1.3, it suffices to show that $\mathcal{F}_{\mathsf{Cap}}(A) = 1$ for all doubly stochastic A. For this, we claim that $\boldsymbol{v} = \mathbf{0}$ achieves the optimal objective value in $\mathcal{F}_{\mathsf{Cap}}(A)$. Plugging this in then yields the desired equality $\mathcal{F}_{\mathsf{Cap}}(A) = 1$.

To see that $\boldsymbol{v} = \mathbf{0}$ is optimal, we just need to verify first-order stationarity, since the objective of Eq. (4) is smooth and convex. Observe that the partial derivative of the objective w.r.t. $\boldsymbol{v}_k$ can be written as

$$-1 + \sum_{i=1}^n \frac{A_{i,k} e^{\boldsymbol{v}_k}}{\sum_{j=1}^n A_{i,j} e^{\boldsymbol{v}_j}}.$$

Plugging in $\boldsymbol{v} = \mathbf{0}$ and using the fact that A is doubly stochastic shows that the gradient is zero and so we're done. □

We postpone the proof of Theorem 1.3 for a moment, as we will eventually give a treatment which also handles mixed volumes. We discuss the volume polynomial next.

2 Deterministically Approximating Mixed Volumes

Let $K_1, \dots, K_m \subseteq \mathbb{R}^n$ be a collection of convex bodies. For two sets $A, B \subseteq \mathbb{R}^n$, write $A + B \stackrel{\mathsf{def}}{=} \{x + y : x \in A, y \in B\}$ for their *Minkowski sum*. Similarly, write $\lambda A \stackrel{\mathsf{def}}{=} \{\lambda x : x \in A\}$ for the *dilation* of A. A well-known theorem of Minkowski says that the function

$$(\lambda_1, \dots, \lambda_m) \mapsto \operatorname{Vol}\left(\sum_{i=1}^m \lambda_i K_i \right) \tag{6}$$

on $\mathbb{R}_{\ge 0}^m$ agrees with a homogeneous polynomial of degree-n over all of $\mathbb{R}^m$. In particular, it may be expressed as

$$\sum_{i_1, \dots, i_n = 1}^m V\left(K_{i_1}, \dots, K_{i_n}\right) \cdot \prod_{j=1}^n \lambda_{i_j},$$

where the coefficients $V\left(K_{i_1}, \dots, K_{i_n}\right)$ are known as the *mixed volume* between the bodies $K_{i_1}, \dots, K_{i_n}$. These coefficients encode a rich set of fundamental geometric quantities such as the surface area and mean width of a convex body. Perhaps surprisingly, they also encode many combinatorial quantities (e.g. linear extensions of a poset, and as it turns out, the permanent of a nonnegative matrix), and have many applications; see e.g. [Sta81; Sta89; Gur10; CP21; CP22]. Finally, the mixed volumes satisfy a number of fascinating properties, including the famous *Alexandrov–Fenchel inequalities*. This is a vast strengthening of the Brunn–Minkowski theory, which says that the volume polynomial Eq. (6) is log-concave.

Given their central importance, a natural question is whether or not we can estimate these coefficients. It turns out we can (at least up to a simply exponential multiplicative factor), using the same technique as what we use for the permanent. The following is the analog of Theorems 1.1 and 1.3 combined into one statement.

Theorem 2.1 ([Gur09]). *Let $\boldsymbol{K} = (K_1, \dots, K_n)$ be a tuple of n convex bodies in $\mathbb{R}^n$, and define*

$$\mathcal{V}_{\mathsf{Cap}}(\boldsymbol{K}) \stackrel{\text{def}}{=} \inf_{\boldsymbol{v}\in\mathbb{R}^n} \left\{ \log \operatorname{Vol}\left(\sum_{i=1}^n e^{\boldsymbol{v}_i} K_i \right) - \sum_{i=1}^n \boldsymbol{v}_i \right\}. \tag{7}$$

Then

$$\log V(\boldsymbol{K}) \le \mathcal{V}_{\mathsf{Cap}}(\boldsymbol{K}) \le \log\left(\frac{n^n}{n!}\right) + \log V(\boldsymbol{K}),$$

and there is a deterministic polynomial-time algorithm for estimating $V(\boldsymbol{K})$ up to a multiplicative factor of $\frac{n^n}{n!} \approx e^n$.

Remark 2. There are some technicalities regarding efficient computation of Eq. (7), although these are essentially resolved by the well-known FPRAS for computing the volume (and surface area) of any convex body. See [Gur09] for more details.

Remark 3. One can show that $V(\boldsymbol{K})$ admits an exact variational formulation similar to Eq. (3) since the mixed volume $V(\boldsymbol{K})$ is invariant under permutations. For instance, one can define a distribution over $[n]^n$ given by $\mu(\sigma) \propto V\left(K_{\sigma(1)}, \dots, K_{\sigma(n)}\right)$, etc.

3 Estimating Coefficients of Log-Concave Polynomials

Computing the permanent and mixed volume are examples of a much more general problem.

Problem 1. *Let p be a multivariate polynomial with nonnegative coefficients. Assuming we can (approximately) evaluate p, when can we estimate the coefficients of p?*

Of course, this question only makes sense if p is given as input not as a list of coefficients, but via other means (e.g. in factorized form like Eq. (4), or via membership oracles like Eq. (7)). We show that log-concavity of p on $\mathbb{R}^n_{\ge 0}$ is enough to approximate coefficients up to a multiplicative error of $\frac{n^n}{n!}$. The following theorem encompasses Theorems 1.3 and 2.1 as special cases, where $p(\boldsymbol{z}) = \prod_{i=1}^n \sum_{j=1}^n A_{i,j}\boldsymbol{z}_j$ and $p(\boldsymbol{z}) = \operatorname{Vol}\left(\sum_{i=1}^n \boldsymbol{z}_i K_i\right)$ respectively.

Theorem 3.1 ([Gur06]). *Let $p(\boldsymbol{z})$ be a degree-n homogeneous polynomial over the variables $\boldsymbol{z}_1, \dots, \boldsymbol{z}_n$, and assume p has nonnegative coefficients. Define the* (Gurvits) capacity[1] *by*

$$\mathsf{Cap}_{\mathbf{1}}(p) \stackrel{\text{def}}{=} \inf_{\boldsymbol{z}>0} \frac{p(\boldsymbol{z})}{\boldsymbol{z}_1 \cdots \boldsymbol{z}_n}. \tag{8}$$

If p and all of its partial derivatives are log-concave over $\mathbb{R}^n_{\ge 0}$, then

$$\partial^{\mathbf{1}} p \le \mathsf{Cap}_{\mathbf{1}}(p) \le \frac{n^n}{n!} \cdot \partial^{\mathbf{1}} p$$

We refer interested readers to [KKO21] for applications to the metric traveling salesperson problem (metric TSP), [GL21b; BLP23] for some applications in combinatorics, [Bur+18] for connections to operator scaling, and [GL21a] for recent developments. Again, by instantiating Theorem 3.1 with $p(\boldsymbol{z}) = \prod_{i=1}^n \sum_{j=1}^n A_{i,j}\boldsymbol{z}_j$ and $p(\boldsymbol{z}) = \operatorname{Vol}\left(\sum_{i=1}^n \boldsymbol{z}_i K_i\right)$ respectively, taking logarithms, and applying the change of variables $\boldsymbol{z}_i = e^{\boldsymbol{v}_i}$ for all $i \in [n]$, we recover Theorems 1.3 and 2.1.

[1] One can also define a Gurvits capacity for each monomial $\boldsymbol{z}^{\boldsymbol{\kappa}} \stackrel{\text{def}}{=} \prod_{i=1}^n \boldsymbol{z}_i^{\boldsymbol{\kappa}_i}$ in the support of p via $\mathsf{Cap}_{\boldsymbol{\kappa}}(p) \stackrel{\text{def}}{=} \inf_{\boldsymbol{z}>0} \frac{p(\boldsymbol{z})}{\boldsymbol{z}^{\boldsymbol{\kappa}}}$.

Proof of Theorem 3.1. The first inequality $\partial^{\mathbf{1}} p \le \mathsf{Cap}_{\mathbf{1}}(p)$ holds simply because $\partial^{\mathbf{1}} p$ is the coefficient of $\boldsymbol{z}_1 \cdots \boldsymbol{z}_n$, and p has nonnegative coefficients. The heart of the matter is the upper bound. Rearranging, this is equivalent to the claim that

$$\sup_{\boldsymbol{z}>0} \left\{ \frac{(\partial^{\mathbf{1}} p) \cdot \boldsymbol{z}_1 \cdots \boldsymbol{z}_n}{p(\boldsymbol{z})} \right\} \ge \frac{n!}{n^n}.$$

One can interpret the left-hand side probabilistically as looking at the probability mass of $S_n \subseteq [n]^n$ w.r.t. the distribution over $[n]^n$ specified by the coefficients of p, maximized over all possible exponential tilts.

Let us now do a simple telescoping trick, which corresponds probabilistically to decomposition as a product of conditional marginals. For each $k = 1, \dots, n$, define the polynomials

$$q_k(\boldsymbol{z}_1, \dots, \boldsymbol{z}_k) \overset{\mathsf{def}}{=} \left(\partial_{\boldsymbol{z}_{k+1}} \cdots \partial_{\boldsymbol{z}_n} p\right)(\boldsymbol{z}_1, \dots, \boldsymbol{z}_k, 0, \dots, 0).$$

Note that differentiation (resp. setting $\boldsymbol{z}_{k+1} = \cdots = \boldsymbol{z}_n = 0$) filters out all monomials $\boldsymbol{z}^{\boldsymbol{\kappa}}$ of p such that $\boldsymbol{\kappa}_j = 0$ (resp. $\boldsymbol{\kappa}_j > 1$) for some $j = k+1, \dots, n$. Furthermore, each q_k is log-concave over $\mathbb{R}^k_{\ge 0}$, and homogeneous of degree-k. Then

$$\sup_{\boldsymbol{z}>0} \left\{ \frac{(\partial^{\mathbf{1}} p) \cdot \boldsymbol{z}_1 \cdots \boldsymbol{z}_n}{p(\boldsymbol{z})} \right\} = \sup_{\boldsymbol{z}>0} \left\{ \prod_{k=1}^{n} \frac{\boldsymbol{z}_k \cdot q_{k-1}(\boldsymbol{z}_1, \dots, \boldsymbol{z}_{k-1})}{q_k(\boldsymbol{z}_1, \dots, \boldsymbol{z}_k)} \right\}.$$

We will prove that for each $k = 1, \dots, n$ and every fixed $\boldsymbol{z}_1, \dots, \boldsymbol{z}_{k-1} > 0$,

$$\sup_{\boldsymbol{z}_k>0} \left\{ \frac{\boldsymbol{z}_k \cdot q_{k-1}(\boldsymbol{z}_1, \dots, \boldsymbol{z}_{k-1})}{q_k(\boldsymbol{z}_1, \dots, \boldsymbol{z}_k)} \right\} \ge \left(\frac{k-1}{k}\right)^{k-1}.$$

Taking a product over all k clearly yields the desired claim. To establish this inequality, note that since we have fixed $\boldsymbol{z}_1, \dots, \boldsymbol{z}_{k-1}$ as constants, we have a univariate log-concave polynomial ℓ of degree-k in the variable $\boldsymbol{z}_k$, and our goal is to show that

$$\sup_{\boldsymbol{z}_k>0} \left\{ \frac{\boldsymbol{z}_k \cdot \ell'(0)}{\ell(\boldsymbol{z}_k)} \right\} \ge \left(\frac{k-1}{k}\right)^{k-1}. \tag{9}$$

For convenience, we establish a lower bound of $1/e$ instead, which is weaker but more convenient.[2] This is established in Lemma 3.2 below. A proof of the more refined bound is provided in Appendix A. □

Lemma 3.2. *Let ℓ be a univariate polynomial with nonnegative coefficients. If ℓ is log-concave over $\mathbb{R}_{\ge 0}$, then*[3]

$$1 \ge \sup_{z>0} \left\{ \frac{z \cdot \ell'(0)}{\ell(z)} \right\} \ge \frac{1}{e}.$$

Proof. Since $z \cdot \ell'(0)$ is the degree-1 monomial in ℓ and all coefficients of ℓ are nonnegative, the first inequality is immediate. For the second, observe that concavity of $\log \ell(z)$ implies that

$$\log \ell(z) \le \log \ell(0) + \frac{\ell'(0)}{\ell(0)} \cdot z, \qquad \forall z > 0.$$

It follows that

$$\log \sup_{z>0} \left\{ \frac{z \cdot \ell'(0)}{\ell(z)} \right\} \ge \sup_{z>0} \left\{ \log\left(\frac{\ell'(0)}{\ell(0)} \cdot z\right) - \frac{\ell'(0)}{\ell(0)} \cdot z \right\} = \log \sup_{x>0} \left\{ x e^{-x} \right\} = -1,$$

which yields the claim after exponentiating both sides. □

[2] The sharper inequality can be proved by using the fact that log-concavity of p is equivalent to concavity of $p^{1/n}$ by n-homogeneity.

[3] One should be slightly careful about what happens when e.g. $\ell(z) = z^k$. The more precise inequality is that $\inf_{z>0} \frac{\ell(z)}{z} \ge \ell'(0) \ge \frac{1}{e} \inf_{z>0} \frac{\ell(z)}{z}$.

4 Conclusion

Perhaps one way to give a unified view of the variational and entropy-based ideas we saw in the past few lectures is to observe that combining both the Gibbs Variational Principle and the Donsker–Varadhan representation of KL-divergence, we have that for any background probability measure μ (e.g. uniform) over some ambient domain Ω (e.g. $\{\pm 1\}^n$), and any Hamiltonian $f : \Omega \to \mathbb{R}$,

$$\log \mathbb{E}_\mu \left[e^f\right] = \inf_{g:\Omega\to\mathbb{R}} \sup_\nu \left\{\mathbb{E}_\nu[f-g] + \log \mathbb{E}_\mu \left[e^g\right]\right\}. \tag{10}$$

Note that we can exchange the order of $\inf_g$ and $\sup_\nu$ by von Neumann's Minimax Theorem. One can restrict the choice of g or ν in this "saddle-point" formulation to make it tractable. This captures almost all of the examples we have seen.

- **Naïve Mean-Field:** $\Omega = \{\pm 1\}^n, \mu = \mathsf{Unif}\{\pm 1\}^n$, and f is given as input. In Eq. (10), we let g be arbitrary, and restrict ν to be a product measure independently of g.
- **Anari–Oveis Gharan–Vinzant:** $\Omega = 2^{[n]}, \mu = \mathsf{Unif}2^{[n]}$, and f restricts the domain to bases of the input matroid (i.e. $f = -\infty \cdot \mathbf{1}_{\mathscr{B}}$). In Eq. (10), we let ν be arbitrary, and restrict g to be a linear form independently of ν.
- **Gurvits Capacity:** $\Omega = [n]^n$, μ is a "natural" log-concave measure (e.g. $\mu = \bigotimes_{i=1}^n \mu_i$ where $\mu_i(j) \propto A_{ij}$ in the case of permanent), and f restricts the domain to permutations (i.e. $f = -\infty \cdot \mathbf{1}_{S_n}$). In Eq. (10), we let ν be arbitrary, and restrict g to a special type of linear form: $g(\sigma) = \sum_{i=1}^n \boldsymbol{v}_{\sigma(i)}$ for some $\boldsymbol{v} \in \mathbb{R}^n$.

Note that in all of these examples, we made a severe restriction on one of g, ν, and let the other be arbitrary.

4.1 Open Problems

We conclude with some open problems.

Problem 2. *Does there exist a deterministic polynomial-time algorithm for estimating* $\mathrm{per}(A)$ *up to* $(1\pm\epsilon)^n$*-multiplicative error for any nonnegative matrix* $A \in \mathbb{R}_{\geq 0}^{n\times n}$ *and any fixed constant* $\epsilon > 0$*? Similarly, does there exist such an algorithm for estimating mixed volumes? For the latter, Gurvits conjectured that no* FPRAS *exists [Gur09].*

The following question was raised in [Ris16].

Problem 3. *For spin systems in "high temperature" (e.g. when Glauber dynamics mixes rapidly), is there an* FPTAS *for computing the partition function based on convex programming?*

References

[BLP23] Petter Brändén, Jonathan Leake, and Igor Pak. "Lower bounds for contingency tables via Lorentzian polynomials". In: *Israel Journal of Mathematics* 253 (2023), pp. 43–90 (cit. on p. 3).

[Bur+18] P. Burgisser, C. Franks, A. Garg, R. Oliveira, M. Walter, and A. Wigderson. "Efficient Algorithms for Tensor Scaling, Quantum Marginals, and Moment Polytopes". In: *2018 IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS)*. Los Alamitos, CA, USA: IEEE Computer Society, Oct. 2018, pp. 883–897. DOI: 10.1109/FOCS.2018.00088 (cit. on p. 3).

[CP21] Swee Hong Chan and Igor Pak. "Log-concave poset inequalities". In: *arXiv preprint arXiv:2110.10740* (2021) (cit. on p. 2).

[CP22] Swee Hong Chan and Igor Pak. "Introduction to the combinatorial atlas". In: *Expositiones Mathematicae* 40.4 (2022), pp. 1014–1048. ISSN: 0723-0869. DOI: https://doi.org/10.1016/j.exmath.2022.08.003 (cit. on p. 2).

[Ego81] G.P Egorychev. "The solution of van der Waerden's problem for permanents". In: *Advances in Mathematics* 42.3 (1981), pp. 299–305. ISSN: 0001-8708. DOI: https://doi.org/10.1016/0001-8708(81)90044-X (cit. on p. 2).

[Fal81] D.I. Falikman. "Proof of the van der Waerden conjecture regarding the permanent of a doubly stochastic matrix". In: *Mathematical Notes of the Academy of Sciences of the USSR* 29 (1981), pp. 475–479 (cit. on p. 2).

[GL21a] Leonid Gurvits and Jonathan Leake. "Capacity Lower Bounds via Productization". In: *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*. STOC 2021. Virtual, Italy: Association for Computing Machinery, 2021, pp. 847–858. ISBN: 9781450380539. DOI: 10.1145/3406325.3451105 (cit. on p. 3).

[GL21b] Leonid Gurvits and Jonathan Leake. "Counting matchings via capacity-preserving operators". In: *Combinatorics, Probability and Computing* 30.6 (2021), pp. 956–981. DOI: 10.1017/S0963548321000122 (cit. on p. 3).

[Gur06] Leonid Gurvits. "Hyperbolic polynomials approach to Van der Waerden/Schrijver-Valiant like conjectures: sharper bounds, simpler proofs and algorithmic applications". In: *STOC*. 2006, pp. 417–426 (cit. on pp. 1–3).

[Gur09] Leonid Gurvits. "A polynomial-time algorithm to approximate the mixed volume within a simply exponential factor". In: *Discrete & Computational Geometry* 41.4 (2009), pp. 533–555 (cit. on pp. 3, 5).

[Gur10] Leonid Gurvits. "On multivariate Newton-like inequalities". In: *Advances in Combinatorial Mathematics*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 61–78 (cit. on p. 2).

[JSV04] Mark Jerrum, Alistair Sinclair, and Eric Vigoda. "A Polynomial-Time Approximation Algorithm for the Permanent of a Matrix with Nonnegative Entries". In: *Journal of the ACM* 51.4 (July 2004), pp. 671–697 (cit. on p. 1).

[KKO21] Anna R. Karlin, Nathan Klein, and Shayan Oveis Gharan. "A (Slightly) Improved Approximation Algorithm for Metric TSP". In: *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*. STOC 2021. Virtual, Italy: Association for Computing Machinery, 2021, pp. 32–45. ISBN: 9781450380539. DOI: 10.1145/3406325.3451009 (cit. on p. 3).

[Ris16] Andrej Risteski. "How to calculate partition functions using convex programming hierarchies: provable bounds for variational methods". In: *29th Annual Conference on Learning Theory*. Ed. by Vitaly Feldman, Alexander Rakhlin, and Ohad Shamir. Vol. 49. Proceedings of Machine Learning Research. Columbia University, New York, New York, USA: PMLR, 23–26 Jun 2016, pp. 1402–1416 (cit. on p. 5).

[Sta81] Richard P Stanley. "Two combinatorial applications of the Aleksandrov-Fenchel inequalities". In: *Journal of Combinatorial Theory, Series A* 31.1 (1981), pp. 56–65. ISSN: 0097-3165. DOI: https://doi.org/10.1016/0097-3165(81)90053-4 (cit. on p. 2).

[Sta89] Richard P. Stanley. "Log-Concave and Unimodal Sequences in Algebra, Combinatorics, and Geometry". In: *Annals of the New York Academy of Sciences*. Graph Theory and Its Applications East and West: Proceedings of the First China-USA International Graph Theory Conference 576 (1 1989), pp. 500–535 (cit. on p. 2).

A Proof of Eq. (9)

In the above proof of Theorem 3.1, we only used log-concavity of p to deduce e^n-multiplicative approximation. However, since p is also n-homogeneous, the function $p^{1/n}$ is also concave. This is proved in Lemma A.1. Knowing this, the next observation is to use concavity of $q_k^{1/k}$ and Lemma A.2 below to establish Eq. (9), with the tighter $\left(\frac{k-1}{k}\right)^{k-1}$ lower bound. We prove each of Lemmas A.1 and A.2 in turn.

Lemma A.1. *Let $p(\boldsymbol{z})$ be a degree-n homogeneous polynomial*[4] *with nonnegative coefficients in m variables $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_m$. Then the following are equivalent:*

- *$p^{1/n}$ is concave over $\mathbb{R}^m_{\geq 0}$.*
- *p is log-concave over $\mathbb{R}^m_{\geq 0}$.*

[4] It is not essential that the function be a polynomial.

- *For every* $\boldsymbol{z} \in \mathbb{R}^m_{\geq 0}$, $\nabla^2 p(\boldsymbol{z})$ *has at most one positive eigenvalue.*
- p *is quasiconcave over* $\mathbb{R}^m_{\geq 0}$*, i.e. the superlevel sets* $L_t \overset{\text{def}}{=} \{\boldsymbol{z} \in \mathbb{R}^m_{\geq 0} : p(\boldsymbol{z}) \geq t\}$ *are convex.*

Proof Sketch. One can show that the first item implies the second item just by comparing $\nabla^2 p^{1/n}(\boldsymbol{z})$ and $\nabla^2 \log p(\boldsymbol{z})$. The second item implies the third since $\nabla^2 \log p(\boldsymbol{z}) = \frac{p(\boldsymbol{z}) \cdot \nabla^2 p(\boldsymbol{z}) - \nabla p(\boldsymbol{z}) \nabla p(\boldsymbol{z})^\top}{p(\boldsymbol{z})^2}$ is negative semidefinite by assumption, and we are subtracting a positive semidefinite rank-1 matrix from $\nabla^2 p(\boldsymbol{z})$. We show that the third item implies concavity of $p^{1/n}$ by taking advantage of homogeneity.

Recall that in the previous lecture on log-concave polynomials, we established that a symmetric matrix $A \in \mathbb{R}^{m \times m}_{\geq 0}$ with nonnegative entries has at most one positive eigenvalue if and only if for every $v \in \mathbb{R}^m_{\geq 0}$ and $x \in \mathbb{R}^m$, $\left(x^\top A x\right) \cdot \left(v^\top A v\right) \leq \left(x^\top A v\right)^2$. In particular, for a fixed $\boldsymbol{z} \in \mathbb{R}^m_{\geq 0}$, letting $x \in \mathbb{R}^m$ be an arbitrary test vector and taking $v = \boldsymbol{z}$ itself, we have by assumption that

$$\left(x^\top \cdot \nabla^2 p(\boldsymbol{z}) \cdot x\right) \cdot \left(\boldsymbol{z}^\top \cdot \nabla^2 p(\boldsymbol{z}) \cdot \boldsymbol{z}\right) \leq \left(x^\top \cdot \nabla^2 p(\boldsymbol{z}) \cdot \boldsymbol{z}\right)^2.$$

By n-homogeneity of p, we have $\nabla^2 p(\boldsymbol{z}) \cdot \boldsymbol{z} = (n-1)\nabla p(\boldsymbol{z})$ and $\boldsymbol{z}^\top \cdot \nabla^2 p(\boldsymbol{z}) \cdot \boldsymbol{z} = n(n-1)p(\boldsymbol{z})$; this is Euler's Homogeneous Function Theorem. Rearranging and using the fact that $x \in \mathbb{R}^m, \boldsymbol{z} \in \mathbb{R}^m_{\geq 0}$ were arbitrary, the above exactly says that $\nabla^2 p^{1/n}(\boldsymbol{z}) \preceq 0$ for all $\boldsymbol{z} \in \mathbb{R}^m_{\geq 0}$, i.e. $p^{1/n}$ is concave over $\mathbb{R}^m_{\geq 0}$.

Now let us establish the equivalence of these with quasiconcavity. Clearly, $p^{1/n}$ being concave implies quasiconcavity. For the converse, let $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^m_{\geq 0}$ and $0 \leq \lambda \leq 1$ be arbitrary. If it were the case that $p(\boldsymbol{x}) = p(\boldsymbol{y}) = t$ for some $t \in \mathbb{R}_{\geq 0}$, then we would immediately obtain

$$p(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y})^{1/n} \geq t^{1/n} = \lambda p(\boldsymbol{x})^{1/n} + (1-\lambda) p(\boldsymbol{y})^{1/n},$$

since $\boldsymbol{x}, \boldsymbol{y} \in L_t$ implies $\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y} \in L_t$ by convexity of L_t. At this point, we haven't used homogeneity yet. To handle the general case, the trick is to first rescale so that the analysis in the special case $p(\boldsymbol{x}) = p(\boldsymbol{y})$ becomes applicable. Write $\alpha = p(\boldsymbol{x})^{-1/n}$ and $\beta = p(\boldsymbol{y})^{-1/n}$. Then $p(\alpha \boldsymbol{x}) = p(\beta \boldsymbol{y}) = 1$, and so by the above analysis, $p^{1/n}$ evaluated at *any* convex combination of $\alpha \boldsymbol{x}, \beta \boldsymbol{y}$ is lower bounded by 1. In particular, with weights $\frac{\lambda/\alpha}{\lambda/\alpha + (1-\lambda)/\beta}$ and $\frac{(1-\lambda)/\beta}{\lambda/\alpha + (1-\lambda)/\beta}$, we see that

$$p\left(\frac{\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}}{\frac{\lambda}{\alpha} + \frac{1-\lambda}{\beta}}\right)^{1/n} \geq 1,$$

which, after applying homogeneity again as well as the definition of α, β, amounts to

$$p(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y})^{1/n} \geq \frac{\lambda}{\alpha} + \frac{1-\lambda}{\beta} = \lambda p(\boldsymbol{x})^{1/n} + (1-\lambda) p(\boldsymbol{y})^{1/n}.$$

□

Lemma A.2 (Sharpening of Lemma 3.2)**.** *Let* ℓ *be a degree-*k *univariate polynomial with nonnegative coefficients. If* $\ell^{1/k}$ *is concave over* $\mathbb{R}_{\geq 0}$*, then*

$$1 \geq \sup_{z>0} \left\{\frac{z \cdot \ell'(0)}{\ell(z)}\right\} \geq \left(\frac{k-1}{k}\right)^{k-1}.$$

Proof. We follow the proof of Lemma 3.2. Again, the upper bound is immediate. For the lower bound, concavity of $\ell^{1/k}$ implies that

$$\ell^{1/k}(z) \leq \ell^{1/k}(0) + \frac{\ell'(0)}{\ell(0)} \cdot \frac{\ell^{1/k}(0)}{k} \cdot z,$$

or equivalently,

$$\frac{1}{\ell(z)} \geq \frac{1}{\ell(0)} \cdot \left(1 + \frac{\ell'(0)}{\ell(0)} \cdot \frac{z}{k}\right)^{-k}.$$

It follows that

$$\sup_{z>0} \left\{\frac{z \cdot \ell'(0)}{\ell(z)}\right\} \geq \sup_{z>0} \left\{\frac{\ell'(0) \cdot z}{\ell(0)} \cdot \left(1 + \frac{\ell'(0)}{\ell(0)} \cdot \frac{z}{k}\right)^{-k}\right\} = \sup_{x>0} \left\{x\left(1 + \frac{x}{k}\right)^{-k}\right\} = \left(\frac{k-1}{k}\right)^{k-1}.$$

□